Quantum Communication Protocol Employing Weak Measurements

Alonso Botero[1,*,†] and Benni Reznik[2,‡]
[1] Center for Particle Physics, University of Texas, Austin, Texas 78712
[2] School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel

(August 1, 1999)

Quantum Communication and Cryptography are based on quantum mechanical correlations described by entangled states. We propose here a communication protocol exploiting similar correlations between two events with a definite time-ordering: a) the outcome of a *weak measurement* on a spin, and b) the outcome of a subsequent ordinary measurement, or post-selection on the spin. In our protocol, the receiver, Alice, first generates a "code" by performing weak measurements on a sample of $N \sim 100$ spins. The sample is given to Bob, who later performs a post-selection by measuring the spin along either the y or z directions. The results of the post-selection define a random string or "key", which he then broadcasts publicaly. Using both her previously generated code and this random key, Alice is able to infer the *direction* chosen by Bob in the post-selection. On the other hand, we show that the sample can be made small enough so that an eavesdrppoer cannot decode this directional information exclusively from the key.

Weak measurements [1–3] are a special class of quantum measurements explored in recent years by Aharonov, Vaidman, and others. In one such measurement of any given observable A, the disturbance caused to the system is minimized at the expense of precision in a single trial. Nevertheless, after a large number trials one can determine statistical averages such as the expectation value of A. The distinctive feature of weak measurements has to do with the observed averages when the measured system is *post-selected*. Such averages, the so-called *weak values*, may lie outside the bounds of the spectrum [2] of A. Moreover, they may may vary with the chosen pre- and post-selected ensemble. Hence, weak values carry non-trivial information about the *choice* of measurement used for post-selection. In accordance with causality, these unusual regularities must therefore be *a priori* undetectable, i.e., "hidden in the noise" [3]. Hence, they can only be extracted *a posteriori*, from the correlations between readings of the measurement and the result of the post-selection.

In the protocol suggested in this Letter, we exploit the fact that such correlations may be used as a means of "encrypted signaling". The receiver, Alice, starts with a *code* corresponding to the outcomes of a series of weak measurements on a large sample of spins. She then hands them to Bob, who is about to depart on a long voyage, together with an important question that he can only answer at some later time. When he is ready to respond, Bob performs an ordinary measurement (post-selection) on these spins. His response corresponds to the choice of spin component he then measures; for instance, a measurement of σ_y to signal a "yes" ("I will marry you"), or a measurement of σ_z to signal a "no". The random sequence of results obtained in either of these measurements plays the role of a *key*, which he then broadcasts publicaly. This is just enough information for Alice to bin her previous readings and extract the message from the weak values. On the other hand, the public key is useless to the ever-jealous Eve. At best, she could try to infer Bob's choice by running a statistical analysis on the sequence of bits in the key. We show, however, that by choosing appropriate parameters, an optimal sample size may be found which is still too small for Eve to obtain a reliable estimate of the relative frequencies in the key.

We begin with the details of the protocol by analyzing the weak measurement scheme. Suppose that Alice prepares a sample of N spin-1/2 particles, in the eigenstate $|x+\rangle$ of σ_x, and she wishes to perform a weak measurement of the spin observable

$$A \equiv \frac{1}{\sqrt{2}}(\sigma_x + \sigma_y)\,, \tag{1}$$

with eigenstates $|a\pm\rangle$. We let p be the pointer variable of the measuring device and $|\phi_i\rangle$ its initial state, with a Gaussian wave function $\phi_i(p) = (2\pi\Delta p^2)^{-1/4} e^{-p^2/4\Delta p^2}$ of uncertainty Δp. The weak measurement may then be described by the usual transformation taking an initial product state $|\Psi_i\rangle \equiv |x+\rangle|\phi_i\rangle$ of the spin and the device, to a final entangled state:

$$|\Psi_i\rangle \rightarrow |\Psi_f\rangle = c_+|a+\rangle|\phi+\rangle + c_-|a-\rangle|\phi-\rangle\,, \tag{2}$$

where $c_\pm = \langle a \pm |x+\rangle$, and $|\phi\pm\rangle$ is $|\phi_i\rangle$ shifted in p by ± 1. In contrast however to an ordinary measurement, Δp is here assumed to be so large that the shifted states $|\phi\pm\rangle$ overlap considerably. Nevertheless, the expectation value of p is still shifted by the usual expectation value $\langle A\rangle$:

$$\langle \Psi_f|p|\Psi_f\rangle = \langle x+|A|x+\rangle = \frac{1}{\sqrt{2}}\,. \tag{3}$$

The idea of a weak measurement is thus to extract such systematic shifts of the means from a large sample of identically prepared spins, ensuring at the same time a minimal disturbance of any individual spin. This disturbance is naturally related to the overlap between $|\phi+\rangle$ and $|\phi-\rangle$.

A convenient measure of disturbance D is given in terms of the Fidelity $\mathcal{F}$ [4] as $D \equiv 1 - \mathcal{F}$, where $\mathcal{F}$ is defined as the probability of obtaining back the initial state after the weak measurement, $\mathcal{F} = ||\langle x + |\Psi_f\rangle||^2$; hence, D may be interpreted as the probability of "flipping" the initial direction of the spin. As one verifies, D involves the overlap factor $\langle \phi + |\phi -\rangle = e^{-1/2\Delta p^2}$:

$$D = \frac{\Delta A^2}{2}(1 - e^{-1/2\Delta p^2}) \simeq \frac{1}{8\Delta p^2}\,. \tag{4}$$

We assume in the approximation that D is small and take $\Delta A^2 = \langle A^2\rangle - \langle A\rangle^2 = 1/2$.

Together then with the weakness condition $D \ll 1$, the resulting distribution for the pointer variable must be sufficiently broad that the spectrum of A cannot be resolved. The distribution therefore takes essentially a Gaussian form, centered at the expectation value $\langle A\rangle$, with a large uncertainty $\sqrt{\Delta p^2 + \Delta A^2} \simeq \Delta p$ of order $D^{-1/2}$. Since in N identical and independent trials the expected sample mean is $\bar{p} = \langle A\rangle \pm \Delta p/\sqrt{N}$, Alice can determine $\langle A\rangle$ if her sample is large enough.

But now suppose that after completing her measurements Alice gives the spins to Bob, who later performs a second, now ordinary measurement on each spin, along some arbitrary direction $\hat{b}$ with possible outcomes $|b\pm\rangle$. Each outcome in this *post-selection* is correlated with a particular response of the measuring device. To keep track of these correlations it is convenient to re-express the combined state of the system and apparatus in terms of the basis $|b\pm\rangle$

$$|\Psi_f\rangle = |b+\rangle|\phi(b+)\rangle + |b-\rangle|\phi(b-)\rangle\,, \tag{5}$$

where the conditional response of the apparatus is described by the unnormalized states $|\phi(b\pm)\rangle = \langle b \pm |\Psi_f\rangle$. For a given result $|b\rangle$,

$$|\phi(b)\rangle = C\left[\frac{1 + A_w(b)}{2}|\phi+\rangle + \frac{1 - A_w(b)}{2}|\phi-\rangle\right]\,, \tag{6}$$

where $C = \langle b|x+\rangle$ and A_w is the *weak value of* A, defined as

$$A_w(b) \equiv \frac{\langle b|A|x+\rangle}{\langle b|x+\rangle}\,. \tag{7}$$

As expected, the weakness condition in this case entails that $P'(b) = \langle \phi(b)|\phi(b)\rangle$, the probability of obtaining $|b\rangle$ in the presence of the weak measurement, should be very close to the unperturbed probability $P(b) = |\langle b|x+\rangle|^2$. Indeed, the deviation from the unperturbed probability scales with D as

$$\frac{\delta P(b)}{P(b)} = -\frac{1 - |A_w(b)|^2}{\Delta A^2}D\,. \tag{8}$$

Similarly, from the large overlap factor we again expect to find a broad Gaussian-like conditional distribution of the pointer with a width that is still approximately equal to Δp. What is interesting is the location of the mean $\langle p\rangle = \langle \phi_f(b)|p|\phi_f(b)\rangle/\langle \phi_f(b)|\phi_f(b)\rangle$. As is easily verified, this conditional shift in the mean is determined for small D by the weak value of A

$$\langle p\rangle = \frac{\mathrm{Re}A_w(b)}{1 + \frac{\delta P(b)}{P(b)}} = \mathrm{Re}A_w(b) + O(D)\,, \tag{9}$$

and hence, will generally differ from the unconditional expectation value $\langle A\rangle$.

What we see therefore is that if Alice starts with a large sample of N identical spins, she obtains a broad distribution of pointer readings with a sample mean $\langle A\rangle \pm \Delta p/\sqrt{N}$. However, were she to know– *for every single spin*–which of the two possible outcomes $|b+\rangle$ and $|b-\rangle$ was obtained in Bob's measurement, then she could divide her readings into two categories, corresponding to two post-selected sub-samples of size $N_\pm \simeq NP(b\pm)$. Her original distribution would then break up as a mixture of two conditional distributions, with sample means $A_w(b+)$ and $A_w(b-)$ within errors of $\Delta p/\sqrt{N_\pm}$. This break-up is captured in the limit $D \to 0$ by a simple sum rule for the expectation value

$$\langle A\rangle = P(b)\,\mathrm{Re}A_w(b) + P(b-)\,\mathrm{Re}A_w(b-)\,, \tag{10}$$

which is easily verified.

A remarkable property exhibited by this sum rule is that the pair of weak values $\{A_w(b), A_w(b-)\}$ will generally vary with the chosen final direction $\hat{b}$, even when the probabilities $\{P(b), P(b-)\}$ remain unchanged under this variation. In particular, suppose $\hat{b}$ lies on the plane orthogonal to $\hat{x}$, where $P(b) = P(b-) = 1/2$. Letting $\hat{b} = \cos(\theta)\hat{z} + \sin(\theta)\hat{y}$, the average, $\langle A\rangle = 1/\sqrt{2}$ of $A = (\sigma_x + \sigma_y)/\sqrt{2}$ in this case breaks up as an equally weighted sum of the two weak values

$$\mathrm{Re}A_w(b\pm) = \frac{1 \pm \sin(\theta)}{\sqrt{2}}\,. \tag{11}$$

It is this feature of different directions $\hat{b}$ on the equiprobability plane leading to different break-ups of the sample mean which is at the basis of our protocol.

For simplicity let us consider a complete cycle at the end of which Alice receives from Bob a single-bit "yes" or "no" message. When Alice prepares her spin sample in the eigenstate $|x+\rangle$, she labels each spin as $i = 1\,, ..N$. From her N weak measurements of $A = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_y)$, she then generates the code by recording a string of real numbers corresponding to the pointer readings $\{p_1, p_2, ..p_N\}$; the sample mean $\bar{p} = \frac{1}{N}\sum_i p_i$ should yield the expectation value $\langle A\rangle \simeq 1/\sqrt{2}$. Next the spins are given to Bob, carefully keeping track of the ordering. Now it's his turn to send the message. If Bob decides to send a "yes", he measures the σ_y component on every single spin; otherwise he measures σ_z to signal a "no'. In either case, from the resulting sequence of outcomes he generates the key: an ordered list of N bits $\{k_1, k_2, \ldots k_n\}$ where, say, $k_i = 1$ and 0 repectively correspond to the outcomes "up" and "down" for the the i-th spin. It is this key which Bob sends back to Alice using an insecure channel. As the key is effectively random by virtue of the fact that $\hat{y}$ and $\hat{z}$ lie on the equal-probability plane, the uncorrelated sequence of "1" and "0"'s will be useless to the eavesdropper Eve who has tapped the insecure channel.

On the other hand, Alice, upon receiving the key, may go back to her code and separate each reading p_i into either of two bins, depending on whether $k_i = 1$ or 0. Finally, she computes the mean values of p for the two bins

$$\bar{p}_1 = \frac{\sum_i p_i k_i}{\sum_i k_i}, \quad \bar{p}_0 = \frac{\sum_i p_i(1-k_i)}{\sum_i (1-k_i)}. \tag{12}$$

She takes these values as estimates of the "true" means within errors of order $\Delta p/\sqrt{N/2}$. Now, if indeed Bob sent a "yes" message, then she should see that $\bar{p}_0 \simeq A_w(y-) = 0$, and an "eccentric" weak value $\bar{p}_1 \simeq A_w(y+) = \sqrt{2}$. Instead, if Bob sent a "no" message, then she will find no significant deviation in either of the two means $\bar{p}_1$ or $\bar{p}_0$; for this case, the real part of $A_w(z\pm)$ coincides with the sample mean of $\langle A\rangle = 1/\sqrt{2}$. Thus, by distinguishing between a non-trivial and a trivial break-up of the sample mean, Alice can decode Bob's single-bit message.

One may object that due to the small disturbance of the weak measurements, the average numbers of "1"'s and "0"'s in the key are no longer identical for the two final measurements of σ_y and σ_z. Indeed, it is easily seen using eq. (8) that while there is no change in the probabilities for the measurement along z, for the measurement along y the relative change in the probabilities is $\delta P(y\pm)/P(y\pm) = \pm 2D$. The difference may then be used by Eve to distinguish between the two messages if the string is sufficiently long. This leads us to compare two critical values N_w and N_f of the sample size, where N_w is the size required by Alice to detect the message by distinguishing the conditional mean values of p, and N_f the size required by Eve to detect the message from a systematic imbalance in the frequencies of "1"s and "0"s within the key. As we will show, for small disturbance, the ratio between these two numbers scales like

$$\frac{N_f}{N_w} \sim \frac{1}{D} \tag{13}$$

Hence, for a sufficiently small D, an optimal value N may be chosen such that $N_f \gg N \geq N_w$, thus ensuring that the message is securely transmitted.

Consider Alice's requirements. The uncertainty in the sub-sample means goes as $\Delta p/\sqrt{N/2} \sim 1/\sqrt{DN}$; since she needs to distinguish with a resolution of order $\sim O(1)$, this requires an optimal sample number scaling as $N_w \sim \frac{1}{D}$. On the other hand, Eve needs to distinguish

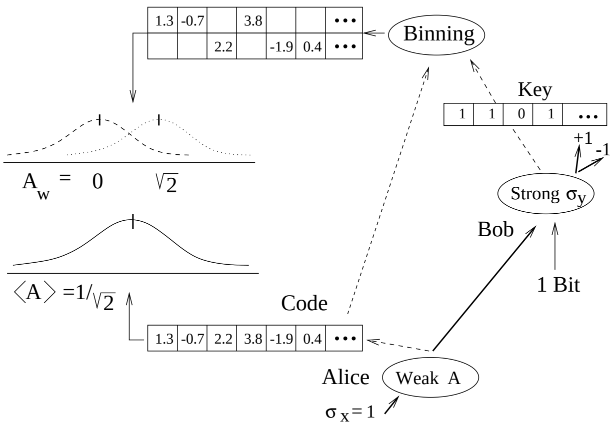

FIG. 1. Schematic of a cycle from which Alice receives a "yes" response from Bob. Upon receiving the key, she bins the readings from her initial weak measurements. Since her original distribution of readings, centered at $1/\sqrt{2}$, breaks up into two distributions centered at the weak values 0 and $\sqrt{2}$, she knows that Bob measured σ_y. A 'no" response would have shown no significant difference from her original distribution.

a change in the frequencies of order δP. She therefore requires that $\delta P > \Delta f$, where the uncertainty in the frequencies Δf for two outcomes and nearly equal probabilities is $P/\sqrt{N}$. Hence, since $\delta P/P$ goes as D, the optimal sample size required by Eve scales as $N_f \sim \frac{1}{D^2}$. We thus verify relation (13). Furthermore, we note that $N_f \simeq (N_w)^2$, which again shows the difference of scales involved in the two methods of extracting Bob's message.

To determine an optimal size, we introduce two relevant parameters: a "secrecy" factor $s \equiv \Delta f/|\delta P(y)|$ indicating the extent to which the imbalance in the frequencies is covered by the noise, and the resolution factor $r \equiv \sqrt{N/2}/\Delta p$, the number of standard deviations (in the sub-sample means) per unit interval. These are given in terms of D and N by $s = 1/2D\sqrt{N}$ and $r \simeq 2\sqrt{DN}$. Fixing r and s and solving for N, we obtain the relation

$$N \simeq s^2 r^4/4\,. \tag{14}$$

For a conservative estimate we take $r = s = 2.7$, yielding an optimal number $N \simeq 100$. We further verify that such settings correspond to the weak regime: the disturbance is $D \simeq 2 \times 10^{-2}$ and from eq. (9) we see that for the $\sigma_y = +1$ post-selection, there is only a 4% correction to the expected ideal mean $\bar{p}_1 = \sqrt{2}$; for $\sigma_y = -1$ and $\sigma_z = \pm 1$ the corresponding means are left unchanged. We anticipate that a more careful examination of the problem is likely to give a lower bound on N.

We would now like to briefly comment on two extensions of the above protocol by which more information is transmitted with larger sample sizes:

First, there is the trivial extension by which Bob transmits to Alice a string of bits $b_1, b_2, ..$, with each bit b_j corresponding to the above procedure performed on a batch of size $\simeq N_w$. Since each bit of information is sent in a separate batch that is independent and uncorrelated from all other batches, Eve cannot extract more information on a single bit of the message by inspecting more batches. The best she could do, when she inspects a public key with a number of bits larger then N_f, is to get an estimate on the average number of "yes" and "no" bits in Bob's message. For any reasonably typical message, this is useless information as she still cannot deduce the order in which they have been sent.

A second extension involves distinguishing m different directions on the equal probability plane. If these are chosen so they lead to equally-spaced weak values, the resolution r must now be scaled by a factor m. However, eq. (14) shows that for large m the method becomes impractical: the optimal number scales as $m^4 r^4$, meaning that the batch size grows exponentially with the number of "directional" bits $\log_2 m$ per batch. It is nevertheless interesting to note that as a matter of principle such directional information can be encoded by this method.

A simple quantum optical realization of the protocol should be feasible following a previous proposal of Aharonov, Davidovich and Zagury [5,6] for observing weak values. These authors considered the interaction of a pre- and post-selected Rydberg atom and the electromagnetic field in a high-Q resonant cavity, where the field is assumed to be initially in a coherent state with a large mean occupation number; this photon number plays the role of the pointer variable. The final post-selection of different atom states corresponds to the measurement direction $\hat{b}$ in our protocol.

To conclude with, it is instructive to compare the present suggestion to the Bennett-Wiesner communication scheme [7] for transmitting 2-bits by means of operations on a single EPR state. In the latter case, Alice prepares an EPR pair and hands (or sends) one of the particles to Bob. When Bob wishes to send a 2-bit message to Alice, he performs a π rotation of his spin around the x, y or z axis, or does nothing. He then sends the spin back to Alice, who can reveal his actions by measuring a joint observable of the spins. This scheme, as well as other well known schemes such as teleportation [8] or quantum key distribution [9], evidently rely on the existence of quantum mechanical entanglement and its preservation [10]. In contrast, while in the present proposal the weak measurement can be formulated in terms of entanglement between the measuring device and the system, as in eq. (5), the order of events is here such that the entanglement is no longer "there" once the reading of the weak measurement has been recorded.

How should we then understand the flow of information in the present case? Bennett and Wiesner suggested after Schumacher, that since in their scheme only one qbit is returned to Alice while two bits of information are transmitted, then " *one bit of information is sent forward in time... while the other bit is sent backwards in time to the EPR source, then forward in time though the untreated particle...*". In our example, because the code is prepared before Bob sends his message, and because no useful information can be extracted from the key, the message is in some sense already "in" the code. It seems therefore that the full one bit of information is sent backwards in time. Yet, it is only with the aid of the post-selected "key" that the message is extracted from the quantum noise. Therefore, no conflict arises between macroscopic causality and this apparently retrocausal flow of information.

We thank Y. Aharonov, M. Byrd, A. Casher, T. Efron, M. Mims, Y. Ne'eman, S. Nussinov, and L. Vaidman for helpful discussions. We also thank the Department of Physics at the University of South Carolina for their kind hospitality and support throughout the preparation of the present article. A. B. acknowledges the support of Fondo Colciencias-BID and World Laboratory. B. R. acknowledges the support from grant 614/95 of the Israel Science Foundation, established by the Israel Academy of Sciences and Humanities.

∗ Also at *Centro Internacional de Física, Edificio Manuel Ancízar, Universidad Nacional, Bogotá,Colombia.*
†*Email: botero@physics.utexas.edu*
‡*Email: reznik@post.tau.ac.il*

---

[1] Y.Aharonov, D. Albert, A. Casher, and L. Vaidman, Phys. Lett. A **124** 199 (1987).

[2] Y. Aharonov, D. Albert, and L. Vaidman, Phys. Rev. Lett. **60** 1351 (1988).

[3] Y. Aharonov and L. Vaidman, Phys. Rev. A, **41**, 11 (1990).

[4] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[5] Y. Aharonov. L. Davidovich, and N. Zagury, Phys. Rev. A **48** 1687 (1993)

[6] N. Zagury and A. F. R. de Toledo Piza, Phy. Rev A, **50**, 2908 (1994).

[7] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[8] C.H. Bennett, G. Brassard, C. Crépeau, R. Josza, A. Peres, and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[9] S. Wiesner, Sigact News **15**(1), 78 (1983). C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175. A.K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[10] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996). C. Bennett, D. DiVincenzo, J.A. Smolin, and W. K. Wootters, Phys. Rev. A. **54**, 3824 (1996).